

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-163497

(43)Date of publication of application : 09.06.1992

(51)Int.Cl.

G10L 3/00
G10L 9/14
// G10L 3/02

(21)Application number : 02-289113

(71)Applicant : RICOH CO LTD

(22)Date of filing : 26.10.1990

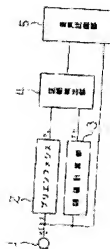
(72)Inventor : MATSUSHITA MITSUGI

(54) VOICE SECTION DETECTING METHOD

(57)Abstract:

PURPOSE: To reduce the occurrence of erroneous detection of a voice section without lowering a processing speed by detecting a section where a voice is present by means of an output through a preemphasis to emphasize the component of a frequency approximately equal to the characteristic frequency region frequency of a vowel having a low voice power.

CONSTITUTION: A voice input part 1 collects voices and converts the voices into an electric signal x_t , and a preemphasis 2 emphasizes the component of a frequency approximately equal to the characteristic frequency region frequency of a voice to increase a voice of a vowel having a low voice power. A threshold calculating part 3 selectively calculates a threshold and calculates a threshold T_h by means of an output signal P_r from the preemphasis 2 at a section where no voice is present. A section detecting part 4 detects a section by means of signals P_r and T_g , an output signal and an input signal therefrom are inputted to a voice input part 5 to recognize a voice.



⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A) 平4-163497

⑬ Int. Cl.⁵ 識別記号 庁内整理番号 ⑭ 公開 平成4年(1992)6月9日
 G 10 L 3/00 3 0 1 A 8842-5H
 9/14 3 0 1 D 8622-5H
 H G 10 L 3/02 3 0 1 8842-5D
 審査請求 未請求 請求項の数 1 (全4頁)

⑮ 発明の名称 音声区間検出方法

⑯ 特 願 平2-289113

⑰ 出 願 平2(1990)10月26日

⑱ 発 明 者 松 下 賢 東京都大田区中馬込1丁目3番6号 株式会社リコー内

⑲ 出 願 人 株 式 会 社 リ コ ー 東京都大田区中馬込1丁目3番6号

⑳ 代 理 人 弁 理 士 柏 木 明

明 細 書

1. 発明の名称 音声区間検出方法

2. 特許請求の範囲

音声を集音して電気信号に変換し、変換された信号を基に音声の存在する区間を検出するようにした音声区間検出方法において、音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するプリエンファスを施した出力を用いて音声の存在する区間を検出するようにしたことを特徴とする音声区間検出方法。

3. 発明の詳細な説明

従来上の利用分野

本発明は、録音下、特に自動車の中、工事場中等などの音声認識用の音声区間検出方法に関する。

従来の技術

音声認識装置を実現する上で、音声区間の検出

は非常に重要であり、以後の認識結果に大きな影響を及ぼす。音声区間検出の一般的なものとしては、例えば「音声認識」(新見孝、共立出版、p.68～69)に示されるように2つの閾値を用いる方法がある。

発明が解決しようとする課題

ところが、一般に母音“い”は他の母音に比べて音声パワーが小さいので、従来法によると欠落が生じ、誤認識してしまうことがある。このような母音“い”の欠落対策として例えば特開昭60-260096号公報に示されるように、区間検出処理前に母音“い”を検出する手段を設け、母音“い”と判断した場合、区間検出のための閾値を下げるという方法があるが、母音“い”を判断する手段を設けなければならず、コストや処理速度の面で問題が残る。

課題を解決するための手段

音声を集音して電気信号に変換し、変換された

特開平4-163497(2)

信号を基に音声の存在する区間を検出するようにした音声区間検出方法において、音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するプリエンファシスを通した出力を用いて音声の存在する区間を検出するようにした。

作用

原音“い”を強調するプリエンファシスをかけた出力信号を用いて音声区間を検出するので、音声パワーの小さい母音“い”の欠落の可能性が低下し、音声区間の誤検出が少なくなる。

実施例

本発明の一実施例を図面に基いて説明する。基本的には、音声入力部1と本実施例の特徴とするプリエンファシス2と隔値計算部3と区間検出部4と音声認識部5よりなる。

このような構成において、本実施例の音声区間検出処理は第2図に示すフローチャートに従い行われる。

ァシス2の出力 P_y とする。

$$Y(f) = w(f) \cdot X(f)$$

$$P_y = \sum_{f=1}^m Y(f)$$

なお、上式中の強調係数 $w(f)$ は母音“い”の第1、2ホルマントなどを強調するもので、発声者が男性の場合は $f=0 \sim 3000\text{Hz}$ （第1ホルマント）、 $1, 0 \sim 2, 4\text{kHz}$ （第2ホルマント）付近で大きな値となる。第4図は強調係数 $w(f)$ の例を示したものである。

また、このようなプリエンファシスの代りに、第4図に示すような特性を持つフィルタなどを用いてもよい。

第5図は、例えば“北見（きみ）”と発声したときの入力信号 $X(t)$ の短時間パワーの時系列 $P_x(n)$ と、プリエンファシスの出力の時系列 $P_y(n)$ とを示している。なお、ここで用いるパワーとは、入力信号の絶対値、或いは自乗値の短

まず、入力音の取込みを行う。音声入力部1はマイクローフのような装置、電気信号変換部によるもので、音声を受信して電気信号 $X(t)$ に変換する。

ついで、プリエンファシス計算処理を行う。プリエンファシス2は母音“い”などのように音声パワーの小さい音声をも大きくするために音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するものであり、例えば第3図に示すように構成されている。このプリエンファシス2は音声パワー検出部6で入力信号 $X(t)$ をバンドパスフィルタ群、或いはFIRなどを施して、 $X(1) \sim X(m)$ なる m 個の帯域に分割した後、 10ms 程度の短時間スペクトル $X(f)$ を抽出し（ただし、 $f=1, 2, \dots, m$ ）、重み付けパワー検出部7で各帯域パワーを次のように重み係数 $w(f)$ に配値してにおいて重み係数 $w(f)$ を用いて、重み付けし、和をとったものをプリエンフ

時間平均値である。

ついで、隔値計算部3による隔値計算を選択的に行う。即ち、音声が存在しない区間におけるプリエンファシス2の出力信号 P_y から隔値 T_h を計算し、記憶する。例えば、

$$T_h = \alpha \cdot P_y + \beta$$

として求める。

一方、区間検出部4によりプリエンファシス2の出力信号 P_y を用いて区間検出を行う。区間検出の方法としては、プリエンファシス2の出力信号 P_y が隔値計算部3で求められた隔値 T_h を一定時間以上越えた区間を音声とする方法、その他の方でもよい。

この区間検出部4の出力信号と入力信号 $X(t)$ とは音声認識部5に入力され、入力信号 $X(t)$ の内で区間検出部4で区間検出された信号のみを用いて音声が認識される。認識方法は、例えば「2値のTSPを用いた単語音声認識システムの例」

特開平 4-163497 (3)

(安田晴則編著、電気学会論文誌C108巻、昭和63年10月号、p858~865)に示されるような音声認識システムでよいが、その他の周知の音声認識システムでもよい。

光明の効果

本発明は、上述したように母音「い」のように音声パワーの小さい母音のホルム・モード周波数近傍の周波数成分を強調するプリエンファシスを通した出力を用いて母音の存在する区間を検出するようにしたので、閉鎖・唇母に音声パワーの小さい母音「い」を含む区間を認識するに際しても、処理速度の遅れを伴うことなく、母音「い」の欠落の可能性を低くできる。音声区間の検出数を減少させることができて、

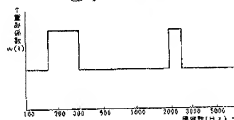
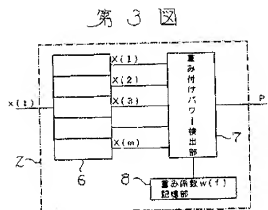
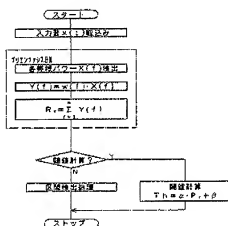
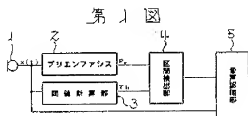
4 図は重み係数群を示す特性図、第 5 図は具体的
なブリエンファシス出力例を示す特性図である。

出 願 人 株式会社 リ コ
代 理 人 柏 木



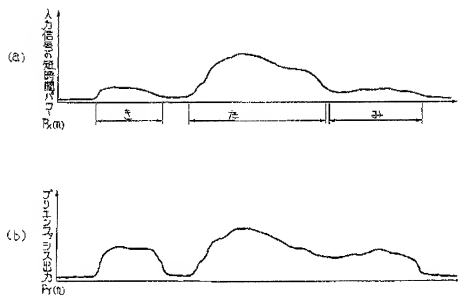
4. 図面の簡単な説明

図面は本発明の一実施例を示すもので、第1図はブロック図、第2図はフローチャート、第3図



特開平 4-163497 (4)

第 5 図



特開平4-163497

